



**HAL**  
open science

# Amélioration de la détection des gouttes dans les systèmes liquide-liquide grâce au couplage entre l'apprentissage automatique et la modélisation assistée par ordinateur

Grégory Bana, Sophie Charton, Fabrice Lamadie, Tojonirina Randriamanantena, Didier Lucor, Nida Sheibat-Othman

## ► To cite this version:

Grégory Bana, Sophie Charton, Fabrice Lamadie, Tojonirina Randriamanantena, Didier Lucor, et al.. Amélioration de la détection des gouttes dans les systèmes liquide-liquide grâce au couplage entre l'apprentissage automatique et la modélisation assistée par ordinateur. CFTL 2024 - Congrès Francophone de Techniques Laser, Sep 2024, Lyon, France. hal-04708954

**HAL Id: hal-04708954**

**<https://hal.science/hal-04708954v1>**

Submitted on 25 Sep 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Amélioration de la détection des gouttes dans les systèmes liquide-liquide grâce au couplage entre l'apprentissage automatique et la modélisation assistée par ordinateur

Grégory BANA<sup>1</sup>, Sophie CHARTON<sup>1</sup>, Fabrice LAMADIE<sup>1\*</sup>, Tojonirina RANDRIAMANANTENA<sup>1</sup>, Didier LUCOR<sup>2</sup>, Nida SHEIBAT-OTHMAN<sup>3</sup>

<sup>1</sup>CEA, DES, ISEC, DMRC, Université de Montpellier, Marcoule, 30200, France

<sup>2</sup>Laboratoire Interdisciplinaire des Sciences du Numérique, Gif-sur-Yvette, 91190, France

<sup>3</sup>Université Claude Bernard Lyon 1, LAGEPP, UMR 5007, CNRS, Villeurbanne, 69100, France

\*(auteur correspondant : : fabrice.lamadie@cea.fr)

**Résumé** - Cette communication présente un nouvel algorithme d'apprentissage automatique pour le traitement d'images destiné à la détection des gouttes dans les systèmes liquide-liquide. Il combine trois outils numériques : YOLOv5 pour la détection d'objets, Blender pour la génération d'images synthétiques, et CycleGAN pour le texturage des images. En conséquence, il a été nommé BYG-Drop pour "Blender-YOLO-cycleGAN Droplet detection". BYG-Drop surpasse les techniques traditionnelles de traitement d'images tant en termes de précision que de nombre de gouttes détectées lors des tests de validation numérique. Appliqué à des images expérimentales, il reste cohérent avec les techniques établies telles que la diffraction laser, tout en surpassant les autres méthodes de traitement d'images. Le texturage des images par CycleGAN permet une adaptation rapide à différents systèmes fluides, augmentant la polyvalence du traitement d'images dans la mesure de la distribution de taille des gouttes. Enfin, le temps de traitement par image est considérablement plus rapide avec cette nouvelle approche.

## 1. Introduction

Les émulsions jouent un rôle significatif dans divers procédés industriels, notamment dans les secteurs de l'alimentation, des produits pharmaceutiques, des cosmétiques et de l'énergie. Dans ces procédés, la caractérisation de la taille des gouttes de phase dispersée est essentielle car elle conditionne au premier ordre la performance [1]. Traditionnellement, la mesure de la taille des gouttes s'appuie sur des algorithmes de traitement d'images généralement classés en deux catégories : les méthodes non paramétriques, également connues sous le nom de méthodes de segmentation, et les méthodes paramétriques, également connues sous le nom de méthodes de reconnaissance de forme. Cette classification dépend du fait qu'elles utilisent ou non des caractéristiques a priori des objets à détecter. Parmi les méthodes paramétriques les plus couramment utilisées, on peut citer la transformée de Hough [2] et ses extensions pour détecter des cercles/disques [3] ou des ellipses [4] ou encore certaines approches dédiées basées sur un prétraitement spécifique de l'image suivi par une recherche de motifs [5]. Les méthodes non paramétriques sont quant à elle plus adaptées aux images binarisées. Elles exploitent des outils dérivés de la morphologie mathématique et des opérations de segmentation, en particulier celles combinant les transformations en distance et la segmentation par lignes de partage des eaux (watershed) [6, 7]. Bien que ces approches soient efficaces, elles restent limitées à la caractérisation des écoulements à phase dispersée diluée et, en cas de fort chevauchement des objets (gouttes, bulles, etc.) dans les images, elles conduisent souvent à de fausses détections.

Dans le domaine de la caractérisation des émulsions, comme dans tous les problèmes de traitement d'images, l'apprentissage profond est en train de remplacer le traitement traditionnel.

C'est particulièrement vrai pour les réseaux neuronaux convolutifs (CNN), qui sont remarquablement efficaces pour la reconnaissance d'objets ou la segmentation d'image via les réseaux Unet [8]. Cependant, pour entraîner de tels réseaux à la détection d'objets, il est crucial de créer une base de plusieurs milliers d'images labellisées ce qui reste un enjeu majeur. Dans ce but, la littérature fait état de plusieurs exemples d'approches basées soit sur des algorithmes déterministes, soit sur une annotation manuelle des images. Par exemple, Patil et al. [9] ont utilisé la transformée de Hough circulaire combinée à un algorithme de filtrage pour détecter des gouttes dans des images en complétant « manuellement » le traitement pour corriger les détections manquées. Cui et al. [10] se sont appuyés sur une annotation manuelle complète, ce qui leur a seulement permis de labelliser une base réduite de 100 images de bulles, juste suffisante pour entraîner un premier masque R-CNN dédié à de la classification. Dans les deux cas, l'annotation manuelle des images reste fastidieuse et est limitée à de petits ensembles de données. L'autre possibilité est de labelliser les images en fonction des résultats obtenus par une autre technique de mesure. Sur cette idée, Pieloth et al. [11] ont entraîné un CNN pour prédire directement les distributions de taille des gouttes (DTG) à partir d'un ensemble de 2500 images, labellisées avec des DTG obtenues par diffraction laser. Dans ce cas se pose la problématique de la représentativité des résultats. A travers ces exemples, on comprend que l'un des enjeux d'une bonne utilisation des approches par Machine Learning est la constitution d'une base d'apprentissage suffisamment grande et correctement labellisée.

Dans cette étude, une nouvelle méthode est proposée et testée pour détecter les gouttes dans les images d'émulsion et mesurer leur distribution de taille. Elle exploite les capacités de deux familles de réseaux. Le premier est un réseau de type CNN de la famille YOLO, You Only Look Once [12, 13, 14], adapté à la détection d'objets. Pour surmonter les limitations imposées par la taille et la labellisation du jeu de données d'apprentissage requis pour entraîner un tel réseau, ce dernier a été créé en utilisant Blender, un logiciel de modélisation 3D qui permet de générer des scènes réalistes. Blender peut facilement créer un grand nombre de scènes 3D géométriquement réalistes, qui sont ensuite transformées en images par projection. Dans une deuxième phase, ces images sont rendues semblables aux images réelles grâce à un transfert de texture réalisé par un second réseau, CycleGAN [15]. La combinaison de ces trois outils numériques fournit une méthode polyvalente et efficace pour détecter les gouttes dans les émulsions, quel que soit le couple fluide étudié. Cette nouvelle méthode a été nommée BYG-Drop, acronyme de Blender-YOLO-cycleGAN Droplet detection.

## 2. Principe de l'algorithme

Le principe de BYG-Drop est résumé dans la Figure 1. Il s'agit d'un processus en trois étapes :

1. Création d'un vaste ensemble de données d'images labellisées de gouttes dont les caractéristiques (taille, position, etc.) sont bien connues à l'aide du logiciel de graphisme 3D Blender.
2. Application de textures à cette base à partir d'images réelles en utilisant le réseau de neurones CycleGAN.
3. Entraînement du réseau détecteur d'objets YOLO sur la base d'images texturées.

Le réseau de neurones utilisé dans cette étude est le réseau de détection d'objets YOLO dans sa version V5. Pour une image donnée, il prédit des cadres de délimitation rectangulaires autour des objets détectés. Ainsi, dans le cas de gouttes sphériques, il fournit des informations sur leur diamètre. Une probabilité de détection est associée à chaque objet, ce qui permet d'ajuster la performance de détection et de ne sélectionner que les objets les plus pertinents. Habituel-

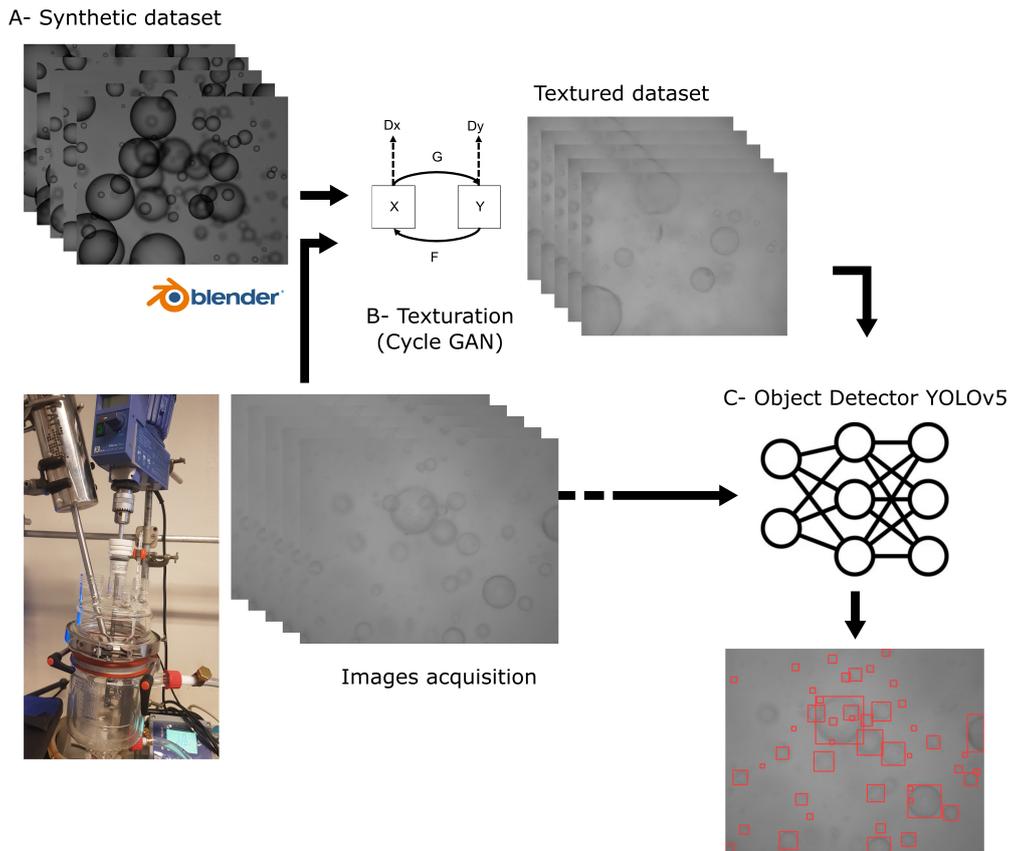


Figure 1 : Principe général de l’algorithme proposé dans cette étude, les images expérimentales sont prises par une sonde endoscopique (SOPAT GmbH) immergée dans un réservoir agité contenant une dispersion eau-huile générée par agitation.

lement, pour entraîner un tel réseau, il est nécessaire de générer une base d’apprentissage de plusieurs milliers d’images labellisées contenant à la fois des objets, ici des gouttes, nets et flous. À cette fin, des images synthétiques ont été générées à partir du logiciel de graphisme 3D Blender. L’utilisation d’images synthétiques permet un contrôle total de la distribution des tailles des gouttes (DTG), assurant ainsi une labellisation efficace. Dans un second temps, et afin de renforcer le processus d’apprentissage, un transfert de texture des images réelles vers les images synthétiques a été réalisé en utilisant CycleGAN. Il permet d’augmenter le réalisme des images synthétisées sans en modifier les propriétés géométriques.

### 3. Résultats et performances

#### 3.1. Validation sur un ensemble d’images synthétiques

Tout d’abord, le réseau a fait l’objet d’une analyse de sensibilité vis-à-vis de deux paramètres : la taille de la base d’apprentissage et l’impact de la texturation par le CycleGAN. Il en ressort que la texturation des images a un impact significatif et est indispensable pour garantir une détection correcte et efficace des gouttes. L’analyse de sensibilité vis-à-vis de la taille de la base d’apprentissage montre quant à elle qu’un minimum de 2000 images est nécessaire pour garantir un entraînement correct du réseau YOLO (chaque image contient en moyenne une cinquantaine d’objets labellisés).

Dans un deuxième temps afin d'évaluer la performance du réseau, un ensemble de 1000 nouvelles images de test a été généré. À des fins de comparaison, des détections ont été effectuées à la fois avec le réseau de neurones et avec la Transformée de Hough Circulaire (THC), en considérant un seuil de confiance de 0,95. La qualité de détection a été évaluée à l'aide de deux métriques classiques du traitement d'image : la précision et le rappel.

- La précision qui mesure le ratio entre le nombre de gouttes correctement détectées et le nombre total de gouttes détectées (combien de gouttes sélectionnées sont pertinentes?).
- Le rappel qui mesure le ratio entre le nombre de gouttes correctement détectées et le nombre total de gouttes correctes (combien de gouttes pertinentes sont sélectionnées?).

Avant d'appliquer la THC, un prétraitement d'image en deux étapes a été effectué pour améliorer le contraste des gouttes par rapport à l'arrière-plan. Il s'agit d'une égalisation d'histogramme uniforme suivie d'une suppression de bruit à l'aide d'un réseau neuronal spécialisé dans le débruitage aveugle [16]. La Figure 2 et le Tableau 1 comparent les résultats obtenus avec les deux algorithmes.

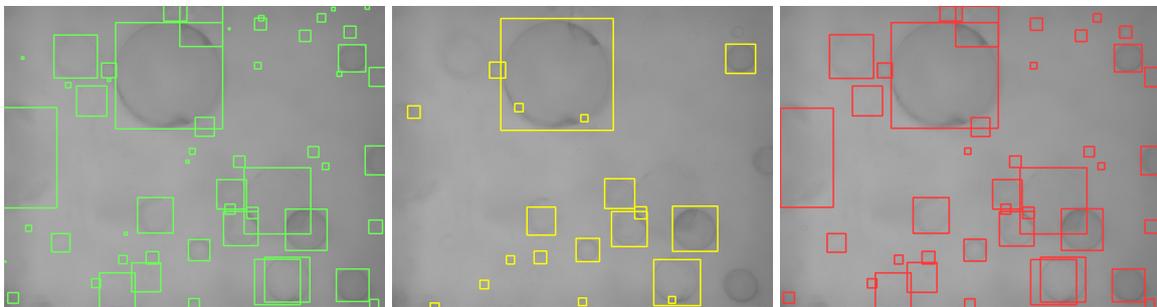


Figure 2 : De gauche à droite, vérité terrain, détections par THC, détections par le nouvel algorithme BYG-Drop.

	BYG-Drop	THC
	-	-
Précision	0,997	0,579
Rappel	0,944	0,265

Tableau 1 : Performances de la THC et de l'algorithme BYG-Drop sur 1000 images synthétiques texturées.

### 3.2. Application expérimentale

Dans un troisième temps, la performance de l'algorithme a été évaluée sur des images réelles par comparaison avec des mesures par granulométrie laser. Le système fluide utilisé comprenait 25 grammes d'huile de silicone (viscosité de  $200 \text{ mP}\cdot\text{s}^{-1}$ ) dispersée dans 975 grammes d'eau ultrapure (viscosité de  $1 \text{ mP}\cdot\text{s}^{-1}$ ), soit une rétention de 2,5%. Trois grammes de Polysorbate (Tween® 20), un surfactant, ont été ajoutés pour stabiliser l'émulsion et empêcher les phénomènes de coalescence pendant le transport vers le granulomètre laser. La vitesse d'agitation dans le réacteur a été ajustée à 600 tr/min et les images ont été acquises grâce à une sonde endoscopique. Cette dernière a été utilisée en mode réflexion avec un réflecteur en téflon blanc et un espace de mesure de 6 mm. La première acquisition d'image a eu lieu 2 minutes après l'ajout

de l'huile (pour assurer l'uniformité dans le réacteur), elle marque le  $t_0$  des expériences. Au total, 11 séries d'acquisition de 200 images chacune, avec une fréquence d'images de 3 images par seconde, ont été réalisées. Les acquisitions d'images ont été espacées de 400 secondes, ce qui donne un temps de mesure total de 67 minutes. Au cours des expériences, quatre échantillons d'environ 75 ml chacun ont été prélevés à :  $t_0+400s$ ,  $t_0+1,200s$ ,  $t_0+2,000s$ , et  $t_0+4,000s$  (fin de l'expérience) pour être analysés avec le granulomètre laser.

Les images ont été traitées avec la THC, l'algorithme BYG-Drop et un troisième algorithme proposé par le fabricant de la sonde (noté SOPAT dans la figure 3). Ce dernier garantit une détection robuste et précise des gouttes en filtrant d'abord la série d'images afin d'éliminer les informations non pertinentes. Pour ce faire, une image moyenne est soustraite de chaque image. Le bruit est ensuite réduit en appliquant un filtre qui normalise l'intensité de chaque pixel en fonction de son environnement local. Pour ce faire, l'image traitée est divisée par une version lissée d'elle-même. La reconnaissance des gouttes est finalement réalisée par un processus de reconnaissance de formes par corrélation avec une base d'images de "bonnes gouttes" préalablement identifiées par l'utilisateur et ayant subi le même prétraitement d'image. La Figure 3 compare les DTG en volume mesurées avec les différentes techniques aux 4 temps de prélèvement. Pour les deux grandes méthodes, imagerie et granulométrie laser, l'évolution des distributions de taille des gouttes est typique de la fragmentation d'une émulsion en réacteur parfaitement agité. Les résultats des deux approches sont en bon accord, sauf pour les plus petites tailles où le nombre de gouttes mesuré par le granulomètre est supérieur à celui obtenu par le traitement d'image, ce qui est un effet connu. Cela est en partie dû à la résolution limitée des images prises par la sonde endoscopique, ce qui rend les plus petits objets indétectables, et d'autre part à la diffraction laser, qui tend à surestimer la fraction des petites gouttes car elles restent plus longtemps dans l'échantillon de mesure (coalescence plus lente).

Les distributions pour les différentes techniques de traitement d'image apparaissent similaires (cf. Figure 3), en partie grâce à la normalisation induite par la présentation de la densité volumique. Cependant, encore une fois, l'approche par apprentissage automatique détecte un nombre de gouttes significativement plus élevé. De plus, les deux autres algorithmes font parfois des erreurs, comme détecter des gouttes fantômes (voir 2 dans la Figure 3d) ou prendre en compte des bulles d'air (voir 1 dans la Figure 3d), ce qui peut biaiser les distributions de taille.

## 4. Conclusion

Dans cette communication, nous avons présenté un nouvel algorithme de traitement d'images par apprentissage automatique, spécialement conçu pour la détection des gouttes dans les systèmes liquide-liquide. Cette méthode novatrice combine trois outils numériques : le détecteur d'objets YOLOv5, des images synthétiques générées par Blender, et un réseau neuronal spécialisé dans le transfert de texture, CycleGAN. En comparaison avec les techniques de traitement d'images traditionnelles, notre approche offre des performances supérieures en termes de précision et de rappel. De plus, elle est cohérente avec les méthodes actuelles de mesure de la distribution de taille des gouttes, telles que la granulométrie laser.

L'utilisation d'images synthétiques pour l'entraînement du réseau, au lieu de recourir à des images réelles annotées manuellement, présente plusieurs avantages majeurs. Elle permet la création de grands ensembles de données, facilitant ainsi un apprentissage efficace. En conséquence, le réseau est capable de minimiser les fausses détections, telles que les gouttes fantômes ou les bulles d'air.

La texturation des images par CycleGAN confère une grande flexibilité, permettant une

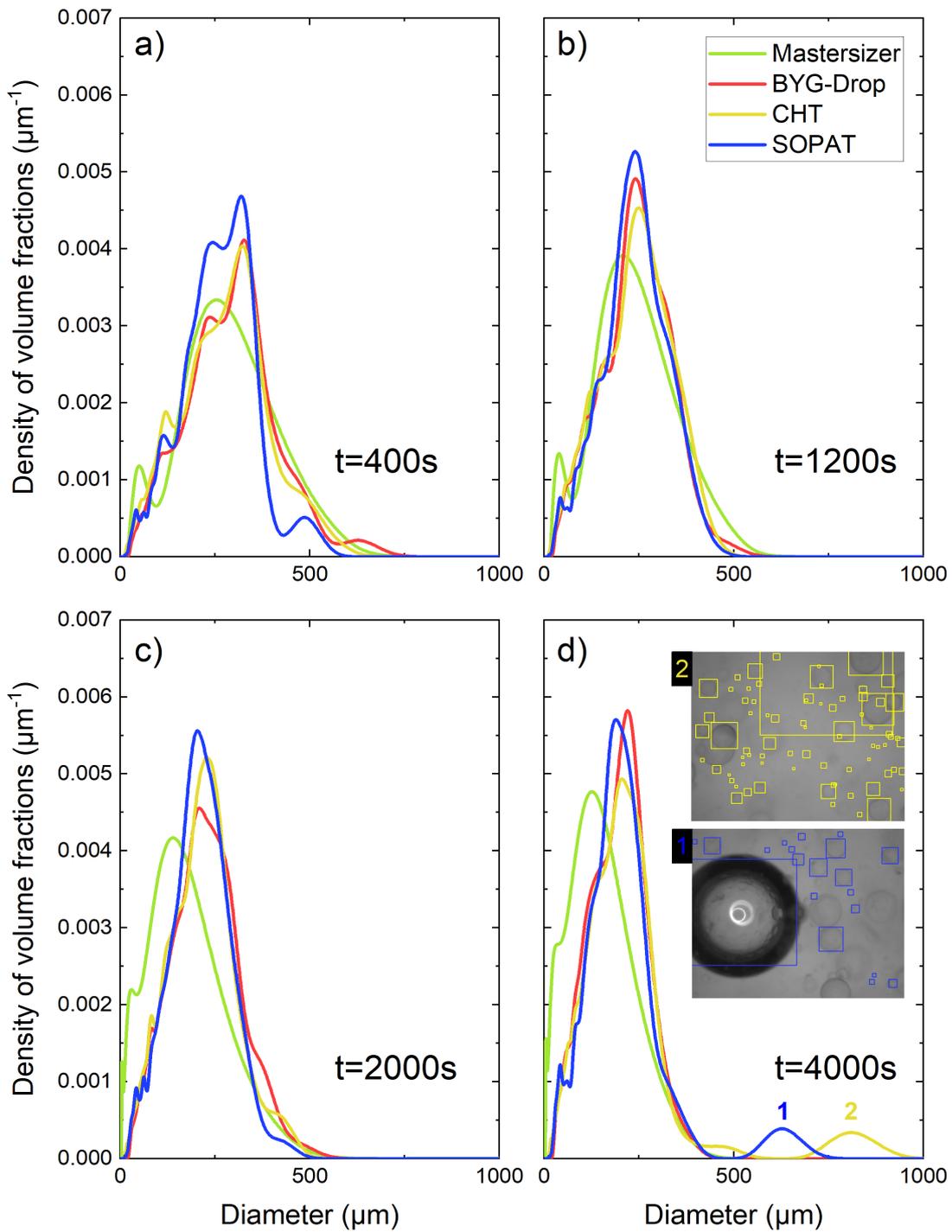


Figure 3 : DTG obtenues à partir du granulomètre à diffraction laser et des trois traitements d'images à  $t_0+400s$ ,  $t_0+1,200s$ ,  $t_0+2,000s$ , et  $t_0+4,000s$ .

adaptation rapide à différents systèmes fluides (colorés ou transparents, gouttes de diverses formes, etc.). Cela renforce la polyvalence de la mesure de la taille des gouttes par imagerie et traitement d'image, la rendant applicable à une large gamme d'environnements.

Enfin, le temps de traitement par image est significativement réduit, grâce à la conception en temps réel de YOLO, même si la version 5 a été utilisée pour un traitement a posteriori dans cette étude. Nous sommes convaincus que BYG-Drop peut être optimisé pour effectuer des mesures en temps réel des distributions de taille des gouttes en utilisant des versions plus avancées du détecteur d'objets et une implémentation sur GPU. L'ensemble des codes et réseaux entraînés sont disponibles sur le dépôt github suivant : <https://github.com/banag0/BYGDrop>.

## Références

- [1] Treybal, FR. E., Mass-Transfer Operations, McGraw-Hill chemical engineering series, 3 (1980)
- [2] Hough, Paul V C, US Patent, A method and means for recognition complex patterns, US Patent : US3069654A (1962)
- [3] Davies, E.R., A modified Hough scheme for general circle location. Pattern Recognition Letters 7, 37–43 (1988)
- [4] Yonghong, Xie and Qiang, Ji, Proceedings - International Conference on Pattern Recognition, 2, A new efficient ellipse detection method, 16 (2002)
- [5] Maaß *et al.*, Automated drop detection using image analysis for online particle size monitoring in multiphase systems, Computers & Chemical Engineering 45 (2012)
- [6] Beucher, S. and Meyer, F., The Morphological Approach to Segmentation : The Watershed Transformation, Mathematical Morphology in Image Processing (2019)
- [7] Soille, P., Morphological Image Analysis : Principles and Applications, Springer-Verlag Berlin Heidelberg GmbH (2004)
- [8] Ronneberger, O. *et al.*, U-Net : Convolutional Networks for Biomedical Image Segmentation, Medical Image Computing and Computer-Assisted Intervention, Springer International Publishing (2015)
- [9] Patil, A., Sægrov, B., Panjwani, B., Advanced deep learning for dynamic emulsion stability measurement. 157 (2022).
- [10] Cui, Y. *et al.*, A deep learning-based image processing method for bubble detection, segmentation, and shape reconstruction in high gas holdup sub-millimeter bubbly flows, Chemical Engineering Journal 449 (2022)
- [11] Pieloth, D. *et al.*, Categorization of Sprays by Image Analysis with Convolutional Neuronal Networks, Chemical Engineering and Technology 46 (2023)
- [12] Redmon, J., Divvala, S., Girshick, R., Farhadi, A., You Only Look Once : Unified, Real-Time Object Detection, in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 779–788 (2016)
- [13] Jiang P. *et al.*, A Review of Yolo Algorithm Developments, Procedia Computer Science 199 (2022)
- [14] Diwan, T. *et al.*, Object detection using YOLO : challenges, architectural successors, datasets and applications, Multimedia Tools and Applications 82 (2023)
- [15] Zhu, J.-Y., Park, T., Isola, P., Efros, A.A., Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks, in 2017 IEEE International Conference on Computer Vision (ICCV), Venice, pp. 2242–2251 (2017).
- [16] Zhang, K. *et al.*, Practical Blind Image Denoising via Swin-Conv-UNet and Data Synthesis, Machine Intelligence Research 20 (2023)

## Remerciements

Ces travaux ont été soutenus par la Direction de l'Énergie du CEA (projet SIACY) et réalisés en partie dans le cadre du Programmes et équipements prioritaires de recherche (PEPR), Recyclage, la Recyclabilité et la Ré-utilisation des matières.